

PM₁₀ 浓度的时间序列模型及预测

程文娜

(四川大学建筑与环境学院,成都 610065)

摘要 利用统计软件 Eviews3.1 对四川省宜宾市 2004 年 7 月 1 日到 8 月 31 日的 PM₁₀ 浓度时间序列数据进行了分析,建立时间序列模型,并对 PM₁₀ 浓度进行预测。结果表明,预测模型精度较高,残差最大值小于 10%,预测结果与实际状况基本相符。

关键词 ARMA 模型 统计分析软件 Eviews3.1 预测可吸入颗粒物或飘尘 PM₁₀

中图法分类号 X830.3; **文献标志码** A

大气环境质量评价是对大气环境状况优劣的定性和定量的评述^[1]。它是认识和研究大气环境质量的一种方法,可以为有效治理和控制大气污染提供必要的科学依据。许多城市随着工业化和商业化的建设,大气环境质量也日趋恶化,这就亟需对城市大气环境质量进行客观、全面、实时的评价。

迄今为止,许多数学方法和计算机技术应用于大气环境质量评价中,如模糊数学^[2]、GIS 技术^[3]、灰色系统^[4]等。文献[5]采用时间序列分析方法,对成都市空气污染指数建立自回归滑动平均模型模拟实测的空气污染指数,并对模拟结果进行了检验。基于这种现实情况,采用时间序列分析的自回归模型对四川省宜宾市 2004 年 7 月 1 日到 8 月 31 日的 PM₁₀ 浓度运用时间序列进行分析建模,并预测 9 月 1 日至 10 日的 PM₁₀ 浓度,分析预测精度。

1 时间序列模型(ARMA)简述

ARMA 模型是 20 世纪 70 年代由计量经济学家 Box-Jenkins 创立的,用于对随机时间序列进行分析研究的一种方法。Box 在文献[6]中对 ARMA(p, q) 模型给出了具体的建模方法。

ARMA 模型把时间序列作为随机过程来研究,必须考虑时间序列的统计特性。虽然对事物变化

的全过程进行一次观测所得的结果,即时间序列的单个变量值的出现具有不确定性,但整个时间序列的变化却具有一定的规律性。一般说来,除少数情况以外,时间序列具有自相关性、动态记忆性等特征。自相关性是指时间序列各观测值之间的依存关系,这种关系可以用相关函数来表示;动态记忆性是从系统的观点来看,在某一时刻进入系统的输入对系统后继行为的影响;自相关性、动态记忆性表现了事物发展的延续性。ARMA 模型就是对时间序列的自相关性、自身的动态记忆性的具体描述,建模的过程就是动态记忆性的量化过程。

根据时间序列动态记忆性的内容不同,ARMA 模型有三种类型:自回归(AR:Auto-Regressive)模型、移动平均(MA:Moving-Average)模型和自回归移动平均(ARMA:Auto-Regressive Moving-Average)模型。

AR(p)即 p 阶自回归模型的形式为

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t。$$

MA(q)即 q 阶移动平均模型的形式为

$$y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}。$$

ARMA(p, q)即 p 阶自回归 q 阶移动平均模型的形式为

$$y_t = \phi_1 y_t + \phi_2 y_t + \cdots + \phi_p y_t + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}。$$

式中, e_t 为随机扰动项。

从以上三种模型的形式可知,AR 模型描述的

是系统对过去自身状态的记忆,MA 模型描述的是系统对过去时刻进入系统的噪声(随机扰动项)的记忆,而 ARMA 模型则是系统对过去自身状态以及进入系统的噪声的记忆。

2 PM₁₀浓度时间序列模型的建立

2004 年 7 月 1 日至 9 月 10 日宜宾市政府的 PM₁₀浓度数据资料见表 1。将此数据序列简记为 y_t 序列,仅对 2004 年 7 月 1 日至 8 月 31 的数据利用 Box-Jenkins 方法建立 ARMA 模型,其过程如下。

表 1 PM₁₀浓度/(mg·m⁻³)

时间	浓度	时间	浓度
2004.07.01	0.092	2004.08.06	0.097
2004.07.02	0.099	2004.08.07	0.101
2004.07.03	0.130	2004.08.08	0.121
2004.07.04	0.158	2004.08.09	0.131
2004.07.05	0.088	2004.08.10	0.131
2004.07.06	0.113	2004.08.11	0.125
2004.07.07	0.139	2004.08.12	0.146
2004.07.08	0.140	2004.08.13	0.101
2004.07.09	0.093	2004.08.14	0.109
2004.07.10	0.100	2004.08.15	0.141
2004.07.11	0.100	2004.08.16	0.09
2004.07.12	0.098	2004.08.17	0.081
2004.07.13	0.118	2004.08.18	0.120
2004.07.14	0.161	2004.08.19	0.152
2004.07.15	0.126	2004.08.20	0.122
2004.07.16	0.129	2004.08.21	0.087
2004.07.17	0.109	2004.08.22	0.120
2004.07.18	0.103	2004.08.23	0.111
2004.07.19	0.131	2004.08.24	0.104
2004.07.20	0.144	2004.08.25	0.094
2004.07.21	0.107	2004.08.26	0.093
2004.07.22	0.116	2004.08.27	0.083
2004.07.23	0.163	2004.08.28	0.094
2004.07.24	0.166	2004.08.29	0.103
2004.07.25	0.169	2004.08.30	0.124
2004.07.26	0.161	2004.08.31	0.117
2004.07.27	0.116	2004.09.01	0.100
2004.07.28	0.103	2004.09.02	0.097
2004.07.29	0.113	2004.09.03	0.115
2004.07.30	0.123	2004.09.04	0.110
2004.07.31	0.099	2004.09.05	0.092
2004.08.01	0.121	2004.09.06	0.092
2004.08.02	0.103	2004.09.07	0.092
2004.08.03	0.102	2004.09.08	0.093
2004.08.04	0.12	2004.09.09	0.115
2004.08.05	0.105	2004.09.10	0.116

数据来源:四川省宜宾市监测站

2.1 PM₁₀浓度时间序列的平稳性检验

利用 Eviews3.1 软件^[7]绘制的 2004 年 7 月 1 日至 8 月 31 日的四川省宜宾市的 PM₁₀浓度时间序列数据图如图 1 所示。从数据图可知该序列没有明显变化趋势。

利用 Eviews3.1 软件对原序列进行 ADF 单位根

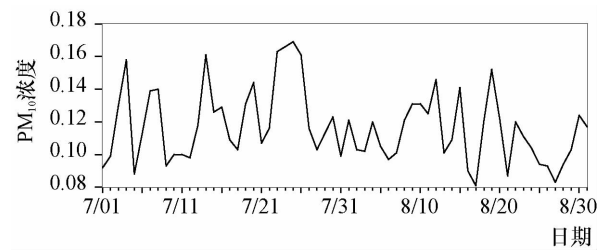


图 1 PM₁₀浓度时间序列图

检验^[8]的结果如表 2。当置信度取 5% 时,ADF 值小于 -2.910 9,可以得知该序列不含单位根,也就是可以认为该序列平稳。序列的自相关和偏自相关系数图也显示大多数自相关和偏自相关系数值落在置信区间里面,因此可以认为该序列是一个随机时间序列。因此,可以对该样本建立 ARMA 模型。

表 2 PM₁₀浓度单位根检验结果

	1% Critical Value	-3.543 7
ADF Test Statistic -3.378 031	5% Critical Value	-2.910 9
	10% Critical Value	-2.592 8

2.2 PM₁₀浓度时间序列模型识别与定阶

对此时间序列利用 Eviews3.1 软件计算自相关和偏自相关函数如图 2。从图中可以初步选定模型

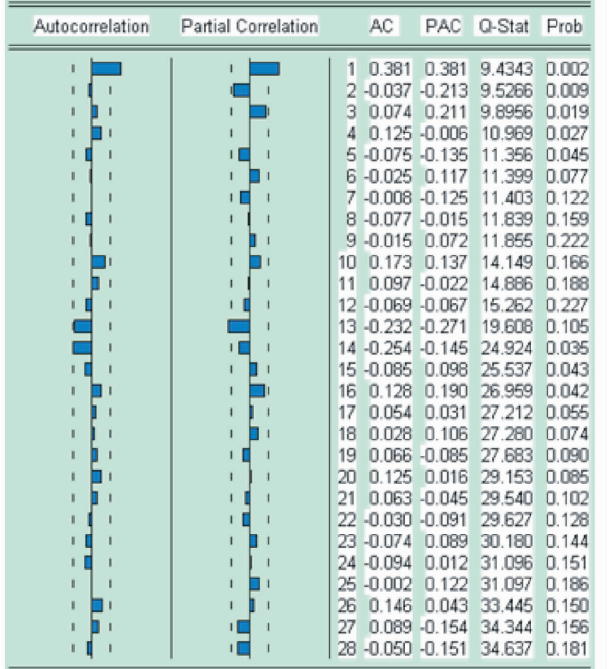


图 2 序列 y_t 的 AC 与 PAC 图

ARMA(1,1), ARMA(2,1), ARMA(3,1), ARMA(5,1), ARMA(7,1), ARMA(13,1), ARMA(14,1), ARMA(13), ARMA(14)。

2.3 PM₁₀ 浓度时间序列模型的建立与检验

对于上述选择的多个阶数不同的 ARMA 模型,运用 Eviews3.1 软件,对序列分别利用以上模型进行分析,通过比较四个模型分析结果,此时间序列 y_t 更适合拟合 AR(14)模型。

对此模型的残差进行检验,得到残差的自相关图和偏自相关图。图形表明,模型的残差都在置信区间范围内,可认为与 0 无明显差异,已基本上消除了自相关和偏自相关,表明残差序列是独立的。

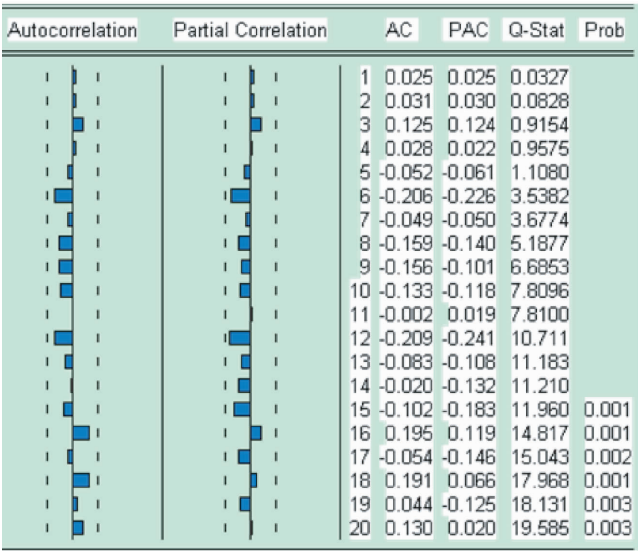


图3 残差序列的 AC 和 PAC 图

综合上述分析可以得出结论,利用 AR(14)模型对 y_t 序列建模是合适的。

3 模型预测

预测是建模的目的之一,预测效果的好坏也是评判模型优劣的标准之一。本文利用 2004 年 7 月 1 日至 8 月 31 日四川省宜宾市委的 PM₁₀ 浓度时间序列数据建立 AR(14)模型,对 2004 年 9 月 1 日至

10 日的 PM₁₀ 浓度进行预测,并与实际值进行比较的结果见图 4,表 4。

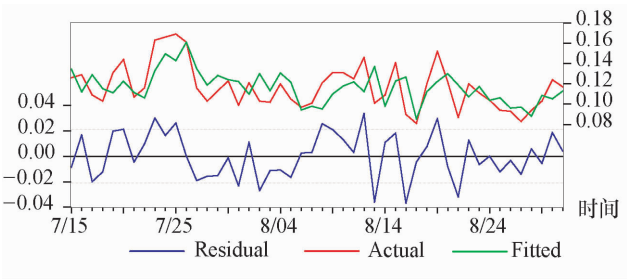


图4 模型拟合效果图

表 4 PM₁₀ 浓度的预测结果/mg · m⁻³

时间	实际值/mg · m ⁻³	预测值/mg · m ⁻³	相对误差/%
2004.09.01	0.100	0.100	0
2004.09.02	0.097	0.096	-1.03
2004.09.03	0.115	0.107	-6.96
2004.09.04	0.110	0.100	-9.09
2004.09.05	0.092	0.092	0
2004.09.06	0.092	0.092	0
2004.09.07	0.092	0.095	3.26
2004.09.08	0.093	0.097	4.30
2004.09.09	0.115	0.105	-8.70
2004.09.10	0.116	0.106	-8.62

结果显示,模型是以 2004 年 7 月 1 日至 8 月 31 日四川省宜宾市委的 PM₁₀ 浓度数据建立的 AR(14)模型,对 2004 年 9 月 1 日至 10 日的 PM₁₀ 浓度进行预测,预测的相对误差分别为:0%; -1.03%; -6.96%; -9.09%; 0%; 0%; 3.26%; 4.30%; -8.70%; -8.62%。预测精度比较高,相对误差在 10% 以内。表明应用时间序列分析预测大气污染物是比较有效的,为大气污染物的预测预报方面提供一定的参考价值。

值得说明的是,时间序列 ARMA 模型所描述的是时间序列的自相关性和自身的动态记忆性,反映的是时间序列的短期变化关系,而不是长期变化关系。因此,利用时间序列模型只能进行短期预测。

(下转第 2266 页)

电价上升,发电商一般都会采取策略性报价来获得更高利润,为抑制发电商利用输电阻塞蓄意抬高市场出清电价必须进行阻塞费用的合理分摊。因此,如何降低阻塞费用,防止投机行为以及由此带来的对发电商报价策略的影响是下一步的研究重点。

参 考 文 献

1 文福拴, David A K. 电力市场中的投标策略. 电力系统自动化, 2000;24(14):1—6
2 杨 根, 周杰娜, 胡志勇. 基于博弈论和概率论的发电商竞价策略研究. 继电器, 2006;34(10):41—44

3 Berry G A, Hobbs B F, Meroney W A, *et al.* Analyzing strategic bidding behavior in transmission networks. In: Singh Hed. IEEE Tutorial on Game Theory Applications in Electric Power Markets. IEEE Power Engineering Society, 1999. 7—32
4 袁智强, 侯志俭, 宋依群, 等. 考虑输电约束古诺模型的均衡分析. 中国电机工程学报, 2004;24(6):73—79
5 余贻鑫, 陈晓明. 考虑输电约束的古诺均衡求解方法. 中国电机工程学报, 2005;25(13):68—72
6 马 莉, 文福拴, 倪以信, 等. 计及网络阻塞影响的发电公司最优报价策略. 电力系统自动化, 2003;27(12):12—17
7 陶 芬, 张步涵, 杨 超. 考虑输电阻塞影响的发电商最优报价策略. 电网技术, 2007;31(16):12—16

The Study of Power Generation Company Bidding Strategy
Considering Transmission Congestion

WU Mei-na, TAN Lun-nong, JIANG Dai-peng
(Jiangsu University, Zhengjiang 212013, P. R. China)

[Abstract] In the electricity market environment, it has become a major concern for generation company of how to build optimal bidding strategies. The problem of developing optimal bidding strategies for generation companies in the electricity market environment with transmission network congestion take into account is systematically investigated, and a two layer optimal model is built under the bidding behaviors of rival generation companies is known, and this model is solved by genetic algorithm. The simulation results of applying the proposed method to IEEE 14-bus test system show that the proposed method can be adopted to guide the drafting of bidding strategy.

[Key words] electricity market bidding strategies congestion management genetic algorithm

(上接第 2262 页)

参 考 文 献

1 酆桂芬. 环境质量评价. 北京: 中国环境科学出版社, 1998
2 季 奎, 戴晓兰. 模糊数学在大气环境质量评价中的应用. 环境科学与管理, 2006;(6):188—190
3 赵 萍, 胡友彪, 桂和荣. 基于 GIS 技术的城市大气环境质量评价——以淮南市为例. 环境科学与技术, 2002;25(4):24—26
4 徐卫国, 田伟利, 张清宇, 等. 灰色关联分析模型在环境空气质量

评价中的修正及应用研究. 中国环境监测, 2006;22(3):66—69
5 柴微涛, 宋述军, 宋学鸿. 成都市城区空气污染指数的时间序列分析. 成都理工大学学报(自然科学版), 2007;34,(4):485—488
6 p. Box G E, Jenkins G M. Time Series Analysis, Forecasting and Control, San Francisco: Halden-Day, 1970
7 张晓峒. 计量经济分析. 北京: 经济科学出版社. 2005:59—60
8 易丹辉. 数据分析与 Eviews 应用. 北京: 中国统计出版社, 2002

Time Series Model and Prediction of PM₁₀ Concentration

CHENG Wen-na

(College of Architecture and Enviroment, Sichuan University, Chengdu 610065, P. R. China)

[Abstract] The PM₁₀ time-series datas of Yibin city, Sichuan Province from July 1, 2004 to October 31, using Eviews3.1 author analyzes. ARMA model is establised and PM₁₀ Concentration is predicted. The result shows that the model has a good prediction effect. Maximum residual is less than 10%. Predictions are consistent with the actual situation of the basic.

[Key words] ARMA model Eviews3.1 prediction PM₁₀